ON THE REDUCTION AND PRESENTATION OF DATA IN ASTRONOMICAL TWO-CHANNEL PHOTOPOLARIMETRY.

G. LEYSHON
Department of Physics and Astronomy,
University of Wales, Cardiff
P.O. Box 913
Cardiff, UK
CF2 3YB

E-mail address: G.Leyshon@astro.cf.ac.uk

ABSTRACT. Many different methods exist for reducing data obtained when an astronomical source is studied with a two-channel polarimeter, such as a Wollaston prism system. This paper presents a rigorous method of reducing the data from raw aperture photometry, and evaluates errors both by a statistical treatment, and by propagating the measured sky noise from each frame. The reduction process performs a hypothesis test for the presence of linear polarization. The probability of there being a non-zero polarization is obtained, and the best method of obtaining the normalized Stokes Parameters is discussed. Point and interval estimates are obtained for the degree of linear polarization, which is subject to positive bias; and the polarization axis is found.

Key words: polarization, methods: data analysis, techniques: polarimetric

1. Introduction

When performing optical polarimetry of astronomical objects, we wish to answer three distinct, but related, physical questions.

Firstly, is the object polarized at all? Secondly, if it is, what is the best estimate of the polarization? And thirdly, what confidence can we give to this measure of polarization?

In addition to these physical questions is a presentational one: in what format should the results be published, so that they will be of most utility to the scientific community?

The questions of quantifying and presenting data on linear polarization have been discussed at length by Simmons & Stewart (1985), who note that the traditional method used by optical astronomers, that of Serkowski (1958), does not give the best estimate of the true polarization under most circumstances. Using their recommendations, I present here a recipe for reducing polarimetric data.

2. Paradigm

In this paper, I will not consider the origin of the polarization of light. It may arise from intrinsic polarization of the source, from interaction with the interstellar medium, or within Earth's atmosphere. Each of these sources represents a genuine polarization, which must be taken into account in explaining the measured polarization values.

Most modern optical polarimetry systems employ a two-channel system, normally a Wollaston prism. Such a prism splits the incoming light into two parallel beams ('channels') with orthogonal polarizations - it functions as a pair of co-located linear analyzers. The transmission axes of the analyzers can be changed either by placing a half-wave plate before the prism in the optical path, and rotating this, or by rotating the actual Wollaston prism.

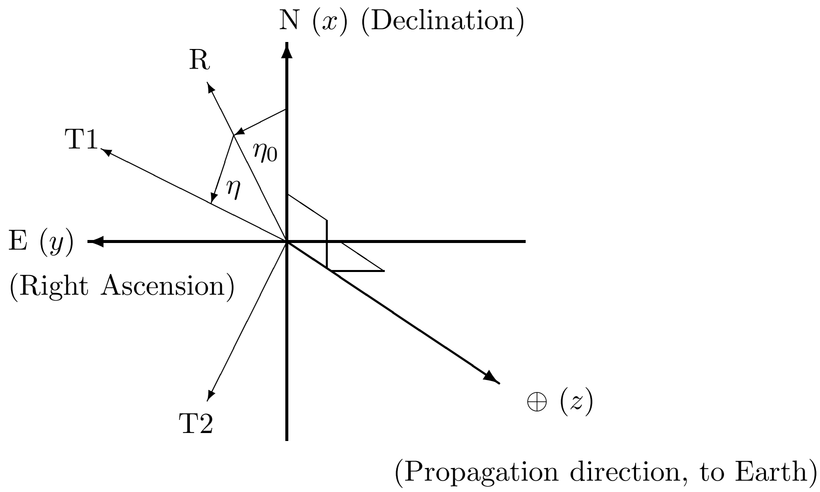

Figure 1. Reference axis, R, relative to celestial co-ordinates.

Such a system is incapable of distinguishing circularly polarized light from unpolarized light, and references to 'unpolarized' light in the remainder of this paper strictly refer to light which is not linearly polarized; it may be totally unpolarized (i.e. randomly polarized), or may include a circularly polarized component.

Where a half-wave plate is used, an anticlockwise rotation χ of the waveplate results in an anticlockwise rotation of $\eta = 2\chi$ of the transmission axes. (For the theory of Wollaston prisms and wave plates, see, for instance, Chapter 8 in Hecht (1987).)

We will suppose that Channel 1 of the detector has a transmission axis which can be rotated by some angle η anticlockwise on the celestial sphere, relative to a reference position η_0 east of north. (See Figure 1.) The transmission axes T1, T2, of Channels 1 and 2 are hence at $\eta_0 + \eta$ and $\eta_0 + 90° + \eta$ respectively.

The reference angle η_0 will depend on the construction of the polarizer, and will not, in general, be neatly due north. For mathematical convenience in the rest of this paper, we will take η_0 to define a reference direction, 'R', in our instrumental co-ordinate system and relate all other angles to it. Such instrumental angles can then be mapped on to the Celestial Sphere by the addition of η_0.

Since the light emerging in the two beams has traversed identical paths until reaching the Wollaston prism, this method of polarimetry does not suffer from the systematic errors due to sky fluctuation which affect single-channel polarimetry (where a single beam polarimeter alternately samples the two orthogonal polarizations).

The two channels will each feed some sort of photometric array, e.g. a CCD, which will record a photon count. Since such images are usually built up by a process of shifting the image position on the array and combining the results, we will refer to a composite image taken in one transmission axis orientation, η, as a *mosaic*. We will denote the rate of arrival of photons recorded in Channel 1 and Channel 2 by $n_1(\eta)$ and $n_2(\eta)$ respectively. From these rates, we can calculate the total intensity (I) of the source, and the difference (S) between the two channels:

$$I(\eta) = n_1(\eta) + n_2(\eta), \tag{1}$$

$$S(\eta) = n_1(\eta) - n_2(\eta). \tag{2}$$

We can also define a *normalized* difference:

$$s(\eta) = \frac{S(\eta)}{I(\eta)}. \tag{3}$$

The purpose of this paper is to discuss how to interpret and present such data.

3. Curve Fitting for p

Suppose we have a beam of light, which has a linearly polarized component of intensity I_p, whose electric vector points at an angle ϕ anticlockwise of R. Its unpolarized component is of intensity I_u. When such a beam enters our detector, we can use Malus' Law (Hecht, 1987, §8.2.1) to deduce that

$$n_1(\eta) = \frac{1}{2} I_u + I_p . \cos^2(\phi - \eta)$$

and

$$n_2(\eta) = \frac{1}{2} I_u + I_p . \sin^2(\phi - \eta),$$

from which we find

$$I(\eta) = I_u + I_p, \tag{4}$$

and, less trivially,

$$S(\eta) = I_p . \cos[2(\phi - \eta)]. \tag{5}$$

The *degree of linear polarization*, p, is defined by

$$p = \frac{I_p}{I_p + I_u} \tag{6}$$

and so we can obtain the normalized difference by substituting Equations 4, 5 and 6 into 3:

$$s(\eta) = p . \cos[2(\phi - \eta)].$$

Now, if observations have been made at a number of different angles, η_j, of the transmission axis, then a series of values for η_j and $s_j(\eta_j)$ will be known, and p and ϕ may be determined by fitting a sine curve to this data, weighted by errors $\sigma_{s_j}(\eta_j)$ as necessary. This method has been used, for example, by di Serego Alighieri *et al.* (1993, §2). (Their refinement of the method allowed for the correction of the $s_j(\eta_j)$ for instrumental polarization at each η_j, which was necessary as they were rotating the entire camera, their system having no half-wave plate.)

We note that if there is any systematic bias of Channel 1 compared to Channel 2, this will show up as an η-independent (DC) term added to the sinusoidal component when $s_j(\eta_j)$ is fitted to the data. Such bias could arise if an object appears close to the edge of the CCD in one channel, for example.

4. The Stokes Parameters

Polarized light is normally quantified using Stokes' parameterisation. (For basic definitions see, for example, Clarke, in Gehrels (ed.) (1974).) Four variables are used, but one, V, is only applicable to circular polarization, which a system involving only half-wave plates and linear analyzers cannot measure. The total intensity, I, of the light is an absolute Stokes Parameter. The other two parameters are defined relative to some reference axis, which in our case will be R, the η_0 direction. Thus we define:

$$Q = S\,(0°) = -S\,(90°),$$

and

$$U = S\,(45°) = -S\,(135°).$$

Normalized Stokes' Parameters are denoted by lower case letters (q,u,v), and are found by dividing the raw parameters by I. We note that S and the normalized s can be thought of as a Stokes Parameter like Q or U, generalised to an arbitrary angle - and results which can be derived for S (or s) will apply to Q and U (or q and u) as special cases.

If the Stokes Parameters are known, then the degree and angle of polarization can be found:

$$p = \sqrt{q^2 + u^2}; \tag{7}$$

$$\phi = \frac{1}{2}.\tan^{-1}(u/q), \tag{8}$$

where the signs of q and u must be inspected to determine the correct quadrant for the inverse tangent. Note that $S\,(\eta)$, Q and U *must* be defined as above to be consistent with the choice of R as Reference.

We must now distinguish between the true values of the Stokes Parameters for a source, and the values which we measure in the presence of noise. We will use the subscript 0 to denote the underlying values, and the subscript i for individual measured values.

In particular, consider a source which is not polarized, so $q_0 = u_0 = 0$, $p_0 = 0$, and ϕ_0 is undefined. Since the q_i and u_i include noise, they will not, in general, be zero, and because of the form of Equation 7, p_i will be a definite-positive quantity. In short, p_i is a *biased* estimator for p_0.

There is no known *unbiased* estimator for p_0, and Simmons & Stewart (1985) discuss at length the question of which estimator should be used. They conclude that the Stokes Parameters themselves are more useful than p and ϕ in many applications, and it is recommended, therefore, that all published polarimetric data should ideally give the normalized Stokes Parameters, with or without evaluation and discussion of p and ϕ.

Given this preference for the Stokes Parameters it appears that one should eschew the curve fitting method in favour of direct evaluation of the parameters, at least when we only have data for the usual angles $\eta_j = 0°, 45°, 90°, 135°$. In practice, observers will take several observations of an object at each transmission angle. This raises the question of how best to combine all the measured values q_i, u_i to yield a single pair of 'best estimators' for q_0 and u_0 – a question which is dealt with by Clarke *et al.* (1983)

On the basis of this prior work and set of recommendations, it is now possible to present a 'recipe' for reducing polarimetric data.

5. Handling the Raw Data

5.1. SKY NOISE AND SHOT NOISE

The raw numbers which our photometric system produces will be a set of photon count rates $n_{1i}(\eta)$ and $n_{2i}(\eta)$, together with their errors, $\sigma_{n_{1i}}(\eta)$ and $\sigma_{n_{2i}}(\eta)$. Errors arise from three sources: photon shot noise; pixel-to-pixel variations in the sky value superimposed on the target object; and imperfect estimation of the modal sky value to subtract from the image (Sterken & Manfroid, 1992; NOAO, IRAF).

The photons arrive at the detector according to a Poisson distribution. Let the total integration time for a mosaic taken at a given rotation of the polarizer, be τ. If the detector requires e photons to arrive in order for one 'count' to be registered, then the total number of photons incident to produce the measured signal is $ne\tau$.

Under Poisson statistics, using units of 'numbers of photons', the standard deviation of the number of photons arriving in this time-bin is the square root of the mean number arriving, *viz.* $\sigma_\gamma = \sqrt{ne\tau}$. In our detector-based count rate units, therefore, the error contributed is $\sigma_{\mathsf{shot}} = \sqrt{n/e\tau}$. Provided $ne\tau > 10$, (Clarke & Cooke, 1983, §19.10) then the shot noise will be normally distributed, to a good approximation.

The modal value of a sky pixel, n_{sky} can be found by considering, say, the pixel values in an annulus of dark sky around the object in question, an annulus which contains $\mathcal{D}$ pixels altogether. The root-mean-square deviation of these pixels' values about the mode can also be found, and we will label this, σ_{sky}. Hence we can estimate the error on the mode, $\sigma_{\mathsf{sky}}/\sqrt{\mathcal{D}}$.

If we perform aperture-limited photometry on our target, with an aperture of area $\mathcal{A}$ in pixels, we must subtract the modal sky level, $\mathcal{A}.n_{\mathsf{sky}}$, which will introduce an error $\sigma_{\mathsf{skysub}} = \mathcal{A}.\sigma_{\mathsf{sky}}/\sqrt{\mathcal{D}}$.

Each individual pixel in the aperture will be subject to a random sky fluctuation; adding these in quadrature for each of the $\mathcal{A}$ pixels, we obtain an error $\sigma_{\mathsf{skyfluc}} = \sqrt{\mathcal{A}}.\sigma_{\mathsf{sky}}$.

Ultimately, the error on the measured, normalized, intensity, is the sum in quadrature of the three quantities, σ_{shot}, σ_{skysub}, and $\sigma_{\mathsf{skyfluc}}$. If the areas of the aperture and annulus are comparable, then both the second and third terms will be significant; in practice, for long exposure times, the first (shot) noise term will be much smaller and can be neglected. This is important as, unlike the sky noise, the shot noise depends on the magnitude of the target object itself. If its contribution to the error terms is negligible, then sky-dominated error terms can be compared between objects of different brightness on the same frame.

DATA CHECK 1. *For each object observed in each channel of each mosaic, the photometry system will have produced a count rate n_i with an error, σ_{n_i}. For each such measurement, calculate $\sqrt{n_i/e\tau}$ and verify that it is much less than σ_{n_i}. Then one can be certain that the noise terms are dominated by sky noise rather than shot noise.*

5.2. TESTING FOR DC BIAS

In practice, for each target object, we will have taken a number of mosaics at each angle η_j. We can immediately use each pair of intensities n_{1i}, n_{2i} to find $I_i(\eta_j)$ and $S_i(\eta_j)$ using Equations 1 and 2.

Since the errors on the two channels are independent, we can trivially find the errors on both $I_i(\eta_j)$ and $S_i(\eta_j)$; the errors turn out to be identical, and are given by:

$$\sigma_{I_i} = \sigma_{S_i} = \sqrt{{\sigma_{n_{1i}}}^2 + {\sigma_{n_{2i}}}^2}. \tag{9}$$

DATA CHECK 2. *Take the mean value of all the $S_i(\eta_j)$ by summing over all the values S_i at all angles η_j; and obtain an error on this mean by combining in quadrature the error on each S_i. If the mean value of $S_i(\eta_j)$, averaged over all the angles η_j, is significantly greater than the propagated error, then there may be some* DC *bias.*

Check 2 uses $S_i(\eta_j)$ as a measure of excess intensity in Channel 1 over Channel 2, and relies on the fact that there are similar numbers of observations at $\eta_j = \eta$ and $\eta_j = \eta + 90°$ to average away effects due to polarization. If, as may happen in real data gathering exercises, there are not *identical* numbers of observations at $\eta_j = \eta$ and $\eta_j = \eta + 90°$, this could show up as apparent 'DC bias' in a highly polarized object. In practice, however, we are unlikely to encounter this combination of events; testing for bias by the above method will either reveal a bias much greater than the error (where the cause should be obvious when the original sky images are examined); or a bias consistent with the random sky noise, in which case we can assume that there is no significant bias.

5.3. OBTAINING THE STOKES PARAMETERS

Once we are satisfied that our raw data are not biased, we can proceed. At this stage in our data reduction, we will find it convenient to divide our set of $S_i(\eta_j)$ values, together with their associated $I_i(\eta_j)$ values, into the named Stokes Parameters,

$$Q_i = S_i(\eta_j = 0°) = -S_i(\eta_j = 90°)$$

and

$$U_i = S_i(\eta_j = 45°) = -S_i(\eta_j = 135°).$$

In the rest of this paper, symbols such as S_i and σ_{S_i}, where not followed by (η), can be read as denoting 'either Q_i or U_i', 'either σ_{Q_i} or σ_{U_i}', etc..

DATA REDUCTION STEP 3. *For each pair of data $n_{1i}(\eta_j), n_{2i}(\eta_j)$, produce the sum, I_i, and the difference, Q_i or U_i as appropriate. Using Equation 9, produce the error common to the sum and difference, σ_{Q_i} or σ_{U_i}. Also find the normalized difference, q_i or u_i.*

In practice, for a given target object, we will have taken a small number of measurements of Q_i and U_i – say ν_Q and ν_U respectively – with individual errors obtained for each measurement. If the errors on the individual values are not comparable, but vary widely, we may need to consider taking a weighted mean.

DATA CHECK 4. *For a set of measurements of (S_i, σ_{S_i}), take all the measured errors, σ_{S_i}; and so find the mean error (call this $\mathcal{E}_{\mathsf{phot}}$) and the maximum deviation of any individual error from $\mathcal{E}_{\mathsf{phot}}$. If the maximum deviation is large compared to the actual error, consider whether you need to weight the data.*

If the deviations are large, we can weight each data point, S_i, by ${\sigma_{S_i}}^{-2}$; but we will not pursue the subject of statistical tests on weighted means here. In practice, one normally finds that the noise does not vary widely between measurements.

We have already checked (see Check 1) that the shot noise is negligible compared with the sky noise terms. Therefore, the main source of variation will be the sky noise. If the maximum deviation of the errors from $\mathcal{E}_{\mathsf{phot}}$ is small, then we can infer that the fluctuation in the sky pixel values is similar in all the mosaics.

DATA REDUCTION STEP 5. *In order to carry the statistical treatment further, we must assume that the sky noise is normally distributed. This is standard astronomical practice.*

DATA REDUCTION STEP 6. *From the sample of Stokes Parameters I_i, Q_i and U_i, obtained in Step 3, find the two means, $\bar{Q}$ and $\bar{U}$, with their corresponding intensities $\bar{I}_Q$ and $\bar{I}_U$; and find the standard deviations of the two* **samples**, *ψ_Q and ψ_U.*

5.4. PHOTOMETRIC AND STATISTICAL ERRORS

Since modern photometric systems can estimate the sky noise on each frame, we are faced throughout our data reduction sequence with a choice between two methods for handling errors. We can propagate the errors on individual measurements through our calculations; or we can use the standard deviation, ψ_S, of the set of sample values, S_i.

In this paper, I use the symbol σ_{S_i} to denote the measured (sky-dominated) error on S_i, and $\sigma_{\bar{S}}$ for the standard error on the estimated mean, $\bar{S}$. The standard deviation of the population, which is the expected error on a single measurement S_i, could be denoted σ_S, but above I used $\mathcal{E}_{\mathsf{phot}}$ to make its photometric derivation obvious.

Using statistical estimators discards the data present in the photometric noise figures and uses only the spread in the data points to estimate the errors. We would expect the statistical estimator to be of similar magnitude to the photometric error in each case; and a cautious approach will embrace the greater of the two errors as the better error to quote in each case.

Because we may be dealing with a small sample (size ν_S) for some Stokes Parameter, S, the standard deviation of the sample, ψ_S, will not be the best estimator of the population standard deviation. The best estimator is (Clarke & Cooke, 1983, §10.5, for example):

$$\mathcal{E}_{\mathsf{stat}} = \sqrt{\frac{\nu_S}{\nu_S - 1}}.\psi_S. \tag{10}$$

In this special case of the **population** standard deviation, I have used the notation $\mathcal{E}_{\mathsf{stat}}$ for clarity. Conventionally, s is used for the 'best estimator' standard deviation, but this symbol is already in use here for a general normalized Stokes Parameter, so in this paper I will use the variant form of sigma, ς, for errors derived from the sample standard deviation, whence $\varsigma_S = \mathcal{E}_{\mathsf{stat}}$, and the (statistical) standard error on the mean is

$$\varsigma_{\bar{S}} = \frac{\psi_S}{\sqrt{\nu_S - 1}} = \frac{\mathcal{E}_{\mathsf{stat}}}{\sqrt{\nu_S}}.$$

The mean value of our Stokes Parameter, $\bar{S}$, is the best estimate of the true value (S_0) regardless of the size of ν_S. Given a choice of errors between $\sigma_{\bar{S}}$ and $\varsigma_{\bar{S}}$, we will cautiously take the greater of the two to be the 'best' error, which we shall denote $\hat{\sigma}_{\bar{S}}$.

DATA CHECK 7. *We now have two ways of estimating the noise on a single measurement of a Stokes Parameter:*

- $\mathcal{E}_{\mathsf{phot}}$ *is the mean sky noise level obtained from our photometry system: Check 4 obtains its value and verifies that the noise levels do not fluctuate greatly about this mean.*
- *Statistical fluctuations in the actual values of the Stokes Parameter in question are quantified by* $\mathcal{E}_{\mathsf{stat}}$, *obtained by applying Equation 10 to the data from Step 6.*

We would expect the two noise figures to be comparable, and this can be checked in our data. We may also consider photometry of other objects on the same frame: Check 1 shows us that the errors are dominated by sky noise, and σ_{sky} *should be comparable between objects, correcting for the different apertures used:*

$$\sigma_{\mathsf{sky}} = \mathcal{E}_{\mathsf{X}}/\sqrt{2\mathcal{A}(1+\mathcal{A}/\mathcal{D})}.$$

We therefore take the best error, $\hat{\sigma}_S$, *on a Stokes Parameter,* S, *to be the greater of* $\mathcal{E}_{\mathsf{phot}}$ *and* $\mathcal{E}_{\mathsf{stat}}$.

If our data passes the above test, then we can be reasonably confident that the statistical tests we will outline in the next sections will not be invalidated by noise fluctuations.

6. Testing for Polarization

The linear polarization of light can be thought of as a vector of length p_0 and phase angle $\theta_0 = 2\phi_0$. There are two independent components to the polarization. If either Q_0 or U_0 is non-zero, the light is said to be polarized. Conversely, if the light is to be described as unpolarized, both Q_0 and U_0 must be shown to be zero.

The simplest way to test whether or not our target object emits polarized light is to test whether the measured Stokes Parameters, $\bar{Q}$ and $\bar{U}$, are consistent with zero. If either parameter is inconsistent with zero, then the source can be said to be polarized.

To proceed, we must rely on our assumption (Step 5) that the sky-dominated noise causes the raw Stokes Parameters, Q_i, U_i, to be distributed normally. Then we can perform hypothesis testing (Clarke & Cooke, 1983, Chapters 12 and 16) for the null hypotheses that Q_0 and U_0 are zero. Here, noting that the number of samples is typically small ($\nu_Q \simeq \nu_U < 30$) we face a choice:

- Either: assume that the sky fluctuations are normally distributed with standard deviation $\mathcal{E}_{\mathsf{phot}}$, and perform hypothesis testing on the standard normal distribution with the statistic:

$$z = \frac{\bar{S} - S_0}{\mathcal{E}_{\mathsf{phot}}/\sqrt{\nu_S}};$$

- Or: use the variation in the S_i values to estimate the population standard deviation $\mathcal{E}_{\mathsf{stat}}$, and perform hypothesis testing on the Student's t distribution with $\nu_S - 1$ degrees of freedom, using the statistic:

$$t = \frac{\bar{S} - S_0}{\mathcal{E}_{\mathsf{stat}}/\sqrt{\nu_S}}.$$

In either case, we can perform the usual statistical test to determine whether we can reject the null hypothesis that '$S_0 = 0$', at the C_S.100% confidence level. The confidence

intervals for retaining the null hypothesis will be symmetrical, and will be of the forms $-z_0 < z < z_0$ and $-t_0 < t < t_0$.

The values of z_0 and t_0 can be obtained from tables, and we define $\bar{S}_{C_S}$ to be the greater of $z_0.\mathcal{E}_{\mathsf{phot}}/\sqrt{\nu_S}$ and $t_0.\mathcal{E}_{\mathsf{stat}}/\sqrt{\nu_S}$. Then the more conservative hypothesis test will reject that null hypothesis at the $C_S.100\%$ confidence level when $|\bar{S}| > \bar{S}_{C_S}$.

In such a confidence test, the probability of making a 'Type I Error', i.e. of identifying an **unpolarized** target as being polarized in *one* polarization sense, is simply $1 - C_S$. The probability of correctly retaining the 'unpolarized' hypothesis is C_S.

The probability of making a 'Type II Error' (Clarke & Cooke, 1983, §12.7) (i.e. not identifying a **polarized** target as being polarized in one polarization sense) is not trivial to calculate.

Now because there are two independent senses of linear polarization, we must consider how to combine the results of tests on the two independent Stokes Parameters. Suppose we have a source which has no linear polarization. We test the two Stokes Parameters, $\bar{Q}$ and $\bar{U}$, for consistency with zero at confidence levels C_Q and C_U respectively. The combined probability of correctly retaining the null hypothesis for both channels is $C_Q.C_U$, and that of making the Type I Error of rejecting the null hypothesis in either or both channels is $1 - C_Q.C_U$. Hence the overall confidence of the combined test is $C_Q.C_U.100\%$.

Since the null hypothesis is that $p_0 = 0$ and ϕ_0 is undefined, there is no preferred direction in the null system, and therefore the confidence test should not prefer one channel over the other. Hence the test must always take place with $C_Q = C_U$.

Even so, the test does not treat all angles equally; the probability of a Type II Error depends on the orientation of the polarization of the source. Clearly if its polarization is closely aligned with a transmission axis, there is a low chance of a polarization consistent with the null hypothesis being recorded on the aligned axis, but a much higher chance of this happening on the perpendicular axis. As the alignment worsens, changing ϕ_0 while keeping p_0 constant, the probabilities for retaining the null hypothesis on the two measurement axes approach one another.

Consider the case where we have taken equal numbers of measurements in the two channels, so $\nu_Q = \nu_U = \nu$, and where the errors on the measurements are all of order $\mathcal{E}_{\mathsf{phot}}$. Hence we can calculate z_0 for the null hypothesis as above. Its value will be common to the Q and U channels, as the noise level and the number of measurements are the same in both channels.

Now suppose that the source has intensity I_0 and a true non-zero polarization p_0 oriented at position angle ϕ_0. Then we can write $Q_0 = I_0 p_0 \cos(2\phi_0)$, and $U_0 = I_0 p_0 \sin(2\phi_0)$. To generate a Type II error, a false null result must be recorded on both axes. The probability of a false null can be calculated for specified p_0 and ϕ_0: defining $z_1 = \frac{I_0 p_0}{\mathcal{E}_{\mathsf{phot}}/\sqrt{\nu}}$ then the probability of such a Type II error is

$$P_{\mathrm{II}} = \frac{1}{2\pi} \int_{x=z_1\cos(2\phi_0)-z_0}^{x=z_1\cos(2\phi_0)+z_0} \int_{y=z_1\sin(2\phi_0)-z_0}^{y=z_1\sin(2\phi_0)+z_0} \exp\left[-\frac{1}{2}(x^2+y^2)\right] dx\, dy. \tag{11}$$

Clearly this probability is not independent of ϕ_0.

DATA REDUCTION STEP 8. *Find the 90% confidence region limits, $\bar{Q}_{90\%}$ and $\bar{U}_{90\%}$, and inspect whether $|\bar{Q}| < \bar{Q}_{90\%}$ and $|\bar{U}| < \bar{U}_{90\%}$.*

- *If both Stokes Parameters fall within the limits, then the target is not shown to be polarized at the 81% confidence level. In this case we can try to find polarization with some*

lower confidence, so repeat the test for $C_Q = C_U = 85\%$. If the null hypothesis can be rejected in either channel, then we have a detection at the 72.25% confidence level. There is probably little merit in plumbing lower confidences than this.

- *If, however, polarization is detected in one or both of the Stokes Parameters at the starting point of 90%, test the polarized parameters to see if the polarization remains at higher confidences, say 95% and 97.5%. The highest confidence with which we can reject the null (unpolarized) hypothesis for either Stokes Parameter should be squared to give the confidence with which we may claim to have detected an overall polarization.*

In our hypothesis testing, we have made the *a priori* assumption that all targets are to be assumed unpolarized until proven otherwise. This is a useful question, as we must ask whether our data are worth processing further − and we ask it using the raw Stokes Parameters, without resorting to complicated formulae. To publish useful results, however, we must produce the normalized Stokes Parameters, together with some sort of error estimate, and it is this matter which we will consider next.

7. The Normalized Stokes Parameters

Consider a general normalized Stokes Parameter for some angle, η:

$$s_i = \frac{S_i}{I_i} = \frac{n_{1i} - n_{2i}}{n_{1i} + n_{2i}}.$$

Clarke *et al.* (1983) point out that the signal/noise ratio obtained by calculating

$$\tilde{s} = \frac{\bar{S}}{\bar{I}} = \frac{\sum_{i=1}^{\nu_S} S_i}{\sum_{i=1}^{\nu_S} I_i} \tag{12}$$

is much better than that obtained by simply taking the mean,

$$\bar{s} = \frac{1}{\nu_s}\sum_{i=1}^{\nu_S} s_i = \frac{1}{\nu_s}\sum_{i=1}^{\nu_S}\frac{S_i}{I_i}, \tag{13}$$

since the Equation 12 involves the taking of only one ratio, where the two terms $\bar{S}$ and $\bar{I}$ have better signal/noise ratios than the individual S_i and I_i which are ratioed in Equation 13.

We also note that errors on $\bar{S}$ and on $\bar{I}$ are not independent of one another. We can write:

$$\tilde{s} = \frac{\bar{n}_1 - \bar{n}_2}{\bar{n}_1 + \bar{n}_2}. \tag{14}$$

If we propagate through the errors on the intensities, we find:

$$\sigma_{\tilde{s}} = \frac{1}{\bar{n}_1 + \bar{n}_2}.\sqrt{[(1-\tilde{s})\sigma_{\bar{n}_1}]^2 + [(1+\tilde{s})\sigma_{\bar{n}_2}]^2}. \tag{15}$$

In order to simplify the calculation, we recall that in Check 4, we checked that the errors on all the S_i (and hence I_i) were similar. Thus the mean error on *one* rate in *one* channel is $\mathcal{E}_{\mathsf{phot}}/\sqrt{2}$. Since the number of measurements made of S is ν_S, then

$$\sigma_{\bar{n}_{1i}} \simeq \sigma_{\bar{n}_{2i}} \simeq \mathcal{E}_{\mathsf{phot}}/\sqrt{2\nu_S}$$

and the error formula approximates to:

$$\sigma_{\tilde{s}} = \tilde{s}.\mathcal{E}_{\mathsf{phot}}.\sqrt{(\bar{S}^{-2} + \bar{I}^{-2})/\nu_S}. \tag{16}$$

In practice, we will be dealing with small polarizations, so $\bar{S} \ll \bar{I}$, and knowing $\tilde{s}$ from Equation 12, then Equation 16 approximates to:

$$\sigma_{\tilde{s}} \simeq \frac{\tilde{s}.\mathcal{E}_{\mathsf{phot}}}{\bar{S}.\sqrt{\nu_S}} = \frac{\mathcal{E}_{\mathsf{phot}}}{\bar{I}.\sqrt{\nu_S}} \tag{17}$$

As we had before with $\mathcal{E}_{\mathsf{stat}}$ and $\mathcal{E}_{\mathsf{phot}}$, so now we have a choice of using sky photometry or the statistics to estimate errors. The above method gives us the photometric error on a normalized Stokes' Parameter as $\varepsilon_{\mathsf{phot}} = \mathcal{E}_{\mathsf{phot}}/\bar{I} = \sigma_{\tilde{s}}.\sqrt{\nu_S}$; the statistical method would be to take the root-mean-square deviation of the measured s_i, obtained in Step 3, about Clarke *et al.*'s (1983) best estimator value, $\tilde{s}$:

$$\varepsilon_{\mathsf{stat}} = \varsigma_{\tilde{s}}.\sqrt{\nu_S} = \frac{1}{\sqrt{\nu_S - 1}}.\left[\sum_{i=1}^{\nu_S}(s_i - \tilde{s})^2\right]^{\frac{1}{2}} \tag{18}$$

DATA REDUCTION STEP 9. *Following the method outlined for finding $\tilde{s}$ and $\sigma_{\tilde{s}}$, apply Equations 12 and 17 to the data obtained in Step 6 to obtain $\tilde{q}$ with $\sigma_{\tilde{q}}$ and $\tilde{u}$ with $\sigma_{\tilde{u}}$.*

DATA CHECK 10. *Using $\tilde{q}$ and $\tilde{u}$, compute $\varsigma_{\tilde{q}}$ and $\varsigma_{\tilde{u}}$; find $\varepsilon_{\mathsf{stat}}$ for both normalized Stokes Parameters, and compare it with $\varepsilon_{\mathsf{phot}}$ in each case. Verify also that the errors, ε_{X}, on the population standard deviations for the two Stokes Parameters are similar – this should follow from the S-independence of Equation 17 for small $\tilde{q}$ and $\tilde{u}$.*

So which error should one publish as the best estimate, $\hat{\sigma}_{\tilde{s}}$, on our final $\tilde{s}$ — $\sigma_{\tilde{s}}$ or $\varsigma_{\tilde{s}}$? Again, a conservative approach would be to take the greater of the two in each case.

DATA REDUCTION STEP 11. *Choose the more conservative error on each normalized Stokes Parameter, and record the results as $\tilde{q} \pm \hat{\sigma}_{\tilde{q}}$ and $\tilde{u} \pm \hat{\sigma}_{\tilde{u}}$. Record also the best population standard deviations, $\hat{\sigma}_q$ and $\hat{\sigma}_u$.*

8. The Degree of Linear Polarization

8.1. THE DISTRIBUTION OF THE NORMALIZED STOKES PARAMETERS

Having obtained estimated values for q and u, with conservative errors, these values – together with the reference angle η_0 – can and should be published as the most convenient form of data for colleagues to work with. It is often desired, however, to express the polarization not in terms of q and u, but of p and ϕ.

Simmons & Stewart (1985) discuss in detail the estimation of the degree of linear polarization. Their treatment assumes that the *normalized* Stokes Parameters have a normal distribution, and that the errors on $\tilde{q}$ and $\tilde{u}$ are similar. This latter condition is true for

small polarizations (see Check 10), but before we can proceed, we must test whether the former condition is satisfied.

If one assumes (Step 5) that n_1 and n_2 are normally distributed, one can construct, following Clarke *et al.* (1983), a joint distribution for s whose parameters are the underlying *population* means (n_{1_0}, n_{2_0}) and standard deviations (σ_1, σ_2) for the photon rates n_{1i} and n_{2i}. The algebra gets a little messy here, so we define three parameters, α, β, γ:

$$\alpha = \frac{1}{2}\left[\frac{1}{{\sigma_1}^2} + \frac{1}{{\sigma_2}^2}\left(\frac{1-s}{1+s}\right)^2\right], \tag{19}$$

$$\beta = \frac{1}{2}\left[\frac{n_{1_0}}{{\sigma_1}^2} + \frac{n_{2_0}}{{\sigma_2}^2}\left(\frac{1-s}{1+s}\right)\right], \tag{20}$$

$$\gamma = \frac{1}{2}\left[\frac{{n_{1_0}}^2}{{\sigma_1}^2} + \frac{{n_{2_0}}^2}{{\sigma_2}^2}\right]. \tag{21}$$

Using these three equations, we can write the probability distribution for s as:

$$P(s) = \frac{\beta.\exp[\frac{\beta^2}{\alpha} - \gamma]}{\sigma_1.\sigma_2.\sqrt{\pi.\alpha^3}.(1+s)^2}. \tag{22}$$

This can be compared to the limiting case of the normal distribution whose mean $\tilde{s}_0$ and standard error σ_0 are obtained by propagating the underlying means (n_{1_0}, n_{2_0}) and standard deviations (σ_1, σ_2) through Equations 14 and 15:

$$P_n(s) = \frac{\exp[\frac{-(s-\tilde{s}_0)^2}{2.{\sigma_0}^2}]}{\sigma_0.\sqrt{2\pi}}; \tag{23}$$

We can derive an expression for the ratio $R(s) = P(s)/P_n(s)$, which should be close to unity if the normalized Stokes Parameter, s, is approximately normally distributed.

DATA CHECK 12.

- *Estimate n_{1_0} and n_{2_0} using Equations 1 and 2, and the data from Step 6. Estimate $\sigma_1 \simeq \sigma_2 \simeq \hat{\sigma}_S/\sqrt{2}$, where $\hat{\sigma}_S$ is obtained from Check 7.*
- *Use the values of $\tilde{s}$ and $\hat{\sigma}_s$ obtained in Step 11 as the best estimates of $\tilde{s}_0$ and σ_0.*
- *Hence use a computer program to calculate and plot $R(s)$ in the domain $-3\hat{\sigma}_s < s < +3\hat{\sigma}_s$. If $R(s)$ is close to unity throughout this domain, then we may treat the normalized Stokes Parameters as being normally distributed.*

8.2. POINT ESTIMATION OF p

If the data passes Checks 10 and 12, then we can follow the method of Simmons & Stewart (1985). They 'normalize' the intensity-normalized Stokes Parameters, q and u, by dividing them by their common population standard deviation, σ. For clarity of notation, in a field where one can be discussing both probability and polarization, I will recast their formulae, such that the *measured* degree of polarization, normalized as required, is here given

in the form $m = \tilde{p}/\sigma$; and the *actual* (underlying) degree of polarization, also normalized, is $a = p_0/\sigma$. It follows from the definition of p (Equation 7) that

$$\sigma_p = \sqrt{\frac{q^2.\sigma_q{}^2 + u^2.\sigma_u{}^2}{q^2 + u^2}}. \tag{24}$$

If $\sigma_q = \sigma_u = \sigma$, then $\sigma_p = \sigma.\sqrt{2}$.

Now, Simmons & Stewart (1985) consider the case of a 'single measurement' of each of q and u, whereas we have found our best estimate of these parameters following the method of Clarke *et al.* (1983) However, we can consider the whole process described by Clarke *et al.* (1983) as 'a measurement', and so the treatment holds when applied to our best estimate of the normalized Stokes Parameters, together with the error on that estimate.

DATA REDUCTION STEP 13. *Find $\hat{\sigma}_p$, and hence $\sigma = \hat{\sigma}_p/\sqrt{2}$, by substituting our best estimates of q and u and their errors (Step 11) ino Equation 24. Hence calculate m:*

$$m = \sqrt{\tilde{q}^2 + \tilde{u}^2}/\sigma.$$

The probability distribution $F(m, a)$ of obtaining a measured value, m, for some underlying value, a, is given by the Rice distribution (Simmons & Stewart, 1985; Wardle & Kronberg, 1974), which is cast in the current notation using the modified Bessel function, I_0 (Boas, 1983, as defined in Ch.12, §17):

$$F(m, a) = m. \exp\left[\frac{-(a^2 + m^2)}{2}\right].I_0(ma) \ldots (m \geq 0) \tag{25}$$

$$F(m, a) = 0 \textit{ otherwise}.$$

Simmons & Stewart (1985) have tested various estimators $\hat{a}_{\mathsf{X}}$ for bias. They find that when $a \lesssim 0.7$, the best estimator is the 'Maximum Likelihood Estimator', $\hat{a}_{\mathsf{ML}}$, which maximises $F(m, a)$ with respect to a. So $\hat{a}_{\mathsf{ML}}$ is the solution for a of:

$$a.I_0(ma) - m.I_1(ma) = 0. \tag{26}$$

If $m < 1.41$ then the solution of this equation is $\hat{a}_{\mathsf{ML}} = 0$.

When $a \gtrsim 0.7$, the best estimator is that traditionally used by radio astronomers, e.g. Wardle & Kronberg (1974). In this case, the best estimator, $\hat{a}_{\mathsf{WK}}$, is that which maximises $F(m, a)$ with respect to m, being the solution for a of:

$$(1 - m^2).I_0(ma) + ma.I_1(ma) = 0. \tag{27}$$

If $m < 1.00$ then the solution of this equation is $\hat{a}_{\mathsf{WK}} = 0$.

Simmons & Stewart (1985) graph $m(a)$ for both cases, and so show that m is a monotonically increasing function of a, and that $\hat{a}_{\mathsf{ML}} < \hat{a}_{\mathsf{WK}} < m \; \forall m$. But which estimator should one use? Under their treatment, the selection of one of these estimators over the other depends on the underlying value of a; they point out that there may be good *a priori* reasons to assume greater or lesser polarizations depending upon the nature of the source.

If we do not make any such assumptions, we can use monotonicity of m and the inequality $\hat{a}_{\mathsf{ML}} < \hat{a}_{\mathsf{WK}} \; \forall m$, to find two limiting cases:

- Let m_{WKmin} be the solution of the Wardle & Kronberg Equation (27) for m with $a = 0.6$. Hence if $m < m_{\mathsf{WKmin}}$, then $\hat{a}_{\mathsf{ML}} < \hat{a}_{\mathsf{WK}} < 0.7$ and the Maximum Likelihood estimator is certainly the most appropriate. Calculating, we find $m_{\mathsf{WKmin}} = 1.0982 \ll 1.41$ and so the Maximum Likelihood estimator will in fact be zero.
- Let m_{MLmax} be the solution of Maximum Likelihood Equation (26) for m with $a = 0.8$. We find $m_{\mathsf{MLmax}} = 1.5347$. Hence if $m > m_{\mathsf{MLmax}}$, then $0.7 < \hat{a}_{\mathsf{ML}} < \hat{a}_{\mathsf{WK}}$, and Wardle & Kronberg's estimator will clearly be the most appropriate.

Between these two extremes, we have $\hat{a}_{\mathsf{ML}} \lesssim 0.7 \lesssim \hat{a}_{\mathsf{WK}}$. This presents a problem, in that each estimator suggests that its estimate is more appropriate than that of the other estimator. If our measured value is $m_{\mathsf{WKmin}} < m < m_{\mathsf{MLmax}}$, what should we take as our best estimate? We could take the mean of the two estimators, but this would divide the codomain of $\hat{a}(m)$ into three discontinuous regions; there might be some possible polarization which this method could never predict! It would be better, then, to interpolate between the two extremes, such that in the range $m_{\mathsf{WKmin}} < m < m_{\mathsf{MLmax}}$,

$$\hat{a} = \frac{m - m_{\mathsf{WKmin}}}{m_{\mathsf{MLmax}} - m_{\mathsf{WKmin}}}.\hat{a}_{\mathsf{ML}} + \frac{m_{\mathsf{MLmax}} - m}{m_{\mathsf{MLmax}} - m_{\mathsf{WKmin}}}.\hat{a}_{\mathsf{WK}}. \tag{28}$$

If we do not know, *a priori*, whether a source is likely to be unpolarized, polarized to less than 1%, or with a greater polarization, then $\hat{a}$ would seem to be a reasonable estimator of the true noise-normalized polarization, and certainly better than the biased m.

DATA REDUCTION STEP 14. *Use the above criteria to find $\hat{a}$, and hence obtain the best estimate, $\hat{p} = \hat{a}.\sigma$, of the true polarization of the target.*

8.3. A CONFIDENCE INTERVAL FOR p

As well as a point estimate for p, we would like error bars. The Rice distribution, Equation 25, gives the probability of obtaining some m given a, and can, therefore, be used to find a confidence interval for the likely values of m given a. We can define two functions, $\mathcal{L}(a)$ and $\mathcal{U}(a)$, which give the lower and upper confidence limits for m, with some confidence C_p; integrating the Rice distribution, these will satisfy:

$$\int_{m=-\infty}^{m=\mathcal{L}(a)} F(m,a).dm = p_1 \tag{29}$$

and

$$\int_{m=\mathcal{U}(a)}^{m=+\infty} F(m,a).dm = p_2 \tag{30}$$

such that

$$1 - C_p = p_1 + p_2. \tag{31}$$

Such confidence intervals are non-unique, and we need to impose an additional constraint. We could require that the tails outside the confidence region be equal, $p_1 = p_2$, but following Simmons & Stewart (1985), we shall require that the confidence interval have the smallest possible width, in which case our additional constraint is:

$$F[\mathcal{U}(a), a] = F[\mathcal{L}(a), a]. \tag{32}$$

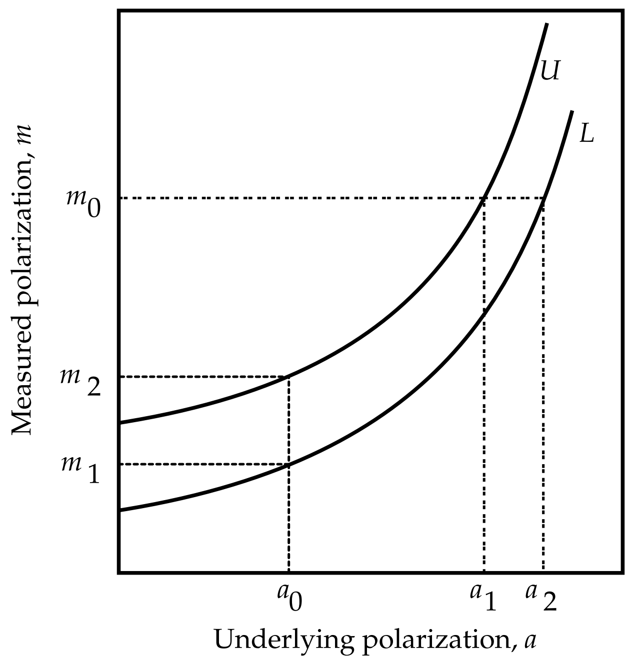

Figure 2. Confidence Intervals based on the Rice Distribution. Figure adapted from Leyshon & Eales (1997).

From the form of the Rice distribution, $\mathcal{L}(a)$ and $\mathcal{U}(a)$ will be monotonically increasing functions of a, as shown in Figure 2. Given a particular underlying polarization a_0, the C_p confidence interval (m_1, m_2) can be obtained by numerically solving Equations 29 thru 32 to yield $m_1 = \mathcal{L}(a_0)$ and $m_2 = \mathcal{U}(a_0)$.

Now, it can be shown (Mood *et al.*, 1974, Ch. VIII, §4.2) that the process can also be inverted, i.e. if we have obtained some measured value m_0, then solving for $m_0 = \mathcal{U}(a_1) = \mathcal{L}(a_2)$ will yield a confidence interval (a_1, a_2), such that the confidence of a lying within this interval is C_p.

Since the contours for $\mathcal{U}(a)$ and $\mathcal{L}(a)$ cut the m-axis at non-zero values of m, we must distinguish three cases, depending on whether or not m_0 lies above one or both of the intercepts. The values of $\mathcal{L}(0)$ and $\mathcal{U}(0)$ depend only on the confidence interval chosen; substituting $a = 0$ into Equations 29 thru 32 results in the pair of equations

$$C_m = \exp\left[-\frac{\mathcal{L}(0)^2}{2}\right] - \exp\left[-\frac{\mathcal{U}(0)^2}{2}\right] \tag{33}$$

and

$$\mathcal{L}(0).\exp\left[-\frac{\mathcal{L}(0)^2}{2}\right] = \mathcal{U}(0).\exp\left[-\frac{\mathcal{U}(0)^2}{2}\right]. \tag{34}$$

A numerical solution of this pair of equations can be found for any given confidence interval, C_m; we find that, in 67% (1σ) interval, $\mathcal{L}(0) = 0.4438$, $\mathcal{U}(0) = 1.6968$, while in a 95% (2σ) interval, $\mathcal{L}(0) = 0.1094$, $\mathcal{U}(0) = 2.5048$. Hence, knowing m_0, and having chosen our desired confidence level, we can determine the interval (a_1, a_2) by the following criteria:

- $\boldsymbol{m_0 \geq \mathcal{U}(0)}$
  .... There are non-zero solutions for both $\mathcal{U}(a_1)$ and $\mathcal{L}(a_2)$.
- $\boldsymbol{\mathcal{L}(0) < m_0 < \mathcal{U}(0)}$
  .... In this case, $a_1 = 0$, and we must solve $m_0 = \mathcal{L}(a_2)$.
- $\boldsymbol{m_0 \leq \mathcal{L}(0)}$
  .... Here, $a_1 = a_2 = 0$.

Simmons & Stewart (1985) note that the third case is formally a confidence interval of zero width, and suggest that this is counter-intuitive; and they go on to suggest an *ad hoc* method of obtaining a non-zero interval. However, it is perfectly reasonable to find a finite probability that the degree of polarization is identically zero: the source may, after all, be unpolarized. This can be used as the basis of estimating the probability that there is a non-zero underlying polarization, as will be shown in the next section.

DATA REDUCTION STEP 15. *Knowing m from Step 13, find the limits (a_1, a_2) appropriate to confidence intervals of 67% and 95%. Hence, multiplying by σ, find the confidence intervals on the estimated degree of polarization. The 67% limits may be quoted as the 'error' on the best estimate.*

8.4. THE PROBABILITY OF THERE BEING POLARIZATION

Consider the contour $m = \mathcal{U}(a)$ on Figure 2. As defined by Equation 30 and the inversion of Mood *et al.* (1974), it divides the domain into two regions, such that there is a probability p_2 of the underlying polarization being greater than $a = \mathcal{U}^{-1}(m_0)$. There is clearly a limiting case where the contour cuts the m-axis at m_0, hence dividing the domain into the polarized region $a > 0$ with probability p_P, and the unpolarized region with probability $1 - p_P$.

Now we may substitute the Rice Distribution, Equation 25, into Equation 30 and evaluate it analytically for the limiting case, $a = 0$:

$$p_P = 1 - \exp(-{m_0}^2/2). \tag{35}$$

Equation 35 hence yields the probability that a measured source actually has an underlying polarization.

DATA REDUCTION STEP 16. *Substitute m from Step 13 into Equation 35. Hence quote the probability that the observed source is truly polarized.*

9. The Polarization Axis

It remains to determine the axis of polarization, for which an unbiased estimate is given by Equation 8. Once again, we have a choice of using the statistical or photometric errors — and, indeed, a choice of raw or normalized Stokes Parameters. Since

$$2\phi = \theta = \tan^{-1}(u/q) = \tan^{-1}(r), \tag{36}$$

our first problem is to obtain the best figure for $r = u/q$.

Now, as we saw in our discussion of the best normalized Stokes Parameter, it is better to ratio a pair of means than to take the mean of a set of ratios. We could take $r = \bar{U}/\bar{Q}$, but for

a very small sample, there is the danger that the mean intensity of the Q observations will differ from that of the U values. Therefore, we should use the normalized Stokes Parameters, and the least error prone estimate of the required ratio will be $\tilde{r} = \tilde{u}/\tilde{q}$, yielding $\tilde{\phi}$.

Knowing the errors on $\tilde{q}$ and $\tilde{u}$, we can find the propagated error in $\tilde{r}$:

$$\sigma_{\tilde{r}} = \tilde{r}.\sqrt{\left(\frac{\tilde{q}}{\sigma_{\tilde{q}}}\right)^2 + \left(\frac{\tilde{u}}{\sigma_{\tilde{u}}}\right)^2}; \tag{37}$$

given the non-linear nature of the tan function, the error on $\tilde{\phi}$ should be found by separately calculating $\sigma_+ = \frac{1}{2}\tan^{-1}(\tilde{r} + \sigma_{\tilde{r}}) - \tilde{\phi}$ and $\sigma_- = \frac{1}{2}\tan^{-1}(\tilde{r} - \sigma_{\tilde{r}}) - \tilde{\phi}$. Careful attention must be paid in the case where the error takes the phase angle across the boundary between the first and fourth quadrants, as the addition of $\pm\pi$ to the inverse tangent may be neccessary to yield a sensible error in the phase angle.

DATA REDUCTION STEP 17. *Obtain $\tilde{\phi}$, the best estimate of ϕ, and the propagated error on it, $\sigma_{\tilde{\phi}} = \frac{1}{2}(|\sigma_+| + |\sigma_-|)$, using Equations 36 and 37. Add η_0 to $\tilde{\phi}$ and hence quote the best estimate of the polarization orientation in true celestial co-ordinates.*

For the statistical error, we note that the probability distribution of observed *phase* angles, $\theta = 2\phi$, calculated by Vinokur (1965), and quoted in Wardle & Kronberg (1974), is:

$$P(\theta) = \exp\left[-\frac{a^2\sin^2(\theta-\theta_0)}{2}\right].$$

$$\left\{\frac{1}{2\pi}\exp\left[-\frac{a^2\cos^2(\theta-\theta_0)}{2}\right] + \frac{a\cos(\theta-\theta_0)}{\sqrt{2\pi}}.\left\{\frac{1}{2} + f[a\cos(\theta-\theta_0)]\right\}\right\} \tag{38}$$

where

$$f(x) = \frac{\mathrm{sign}(x)}{\sqrt{2\pi}}\int_0^x \exp\left(-\frac{z^2}{2}\right)\,dz = \mathrm{sign}(x).\mathrm{erf}(x)/\sqrt{8}, \tag{39}$$

and $\mathrm{erf}(x)$ is the error function as defined in Boas (1983), Ch.11, §9. We do not know $a = p_0/\sigma$, and will have to use our best estimate, $\hat{a}$, as obtained from Step 14. The C_ϕ.100% confidence interval on the measured angle, (θ_1, θ_2), is given by numerically solving

$$\int_{\theta_1}^{\theta_2} P(\theta).d\theta = C_\phi; \tag{40}$$

in this case we choose the symmetric interval, $\theta_2 - \tilde{\theta} = \tilde{\theta} - \theta_1$.

DATA REDUCTION STEP 18. *Obtain the limiting values of $\phi = \theta/2$ for confidence intervals of 67% (1σ) and 95% (2σ). Quote the 67% limits as $\varsigma_{\tilde{\phi}} = (\phi_2 - \phi_1)/2$. Choose the more conservative error from $\varsigma_{\tilde{\phi}}$ and $\sigma_{\tilde{\phi}}$ as the best error, $\hat{\sigma}_{\tilde{\phi}}$.*

10. Comparison with Other Common Techniques

It may be instructive to note how the process of reducing polarimetric data outlined in this paper compares with the methods commonly used in the existing literature. The paper by

Simmons & Stewart (1985) gives a thorough review of five possible point estimators for the degree of polarisation. One of these methods is the trivial m as an estimator of a. The other four methods all involve the calculation of thresholds $\underline{m}_{\mathsf{X}}$: if $m < \underline{m}_{\mathsf{X}}$ then $\hat{a}_{\mathsf{X}} = 0$. These four methods are the following:

1. Maximum Likelihood: as defined above, $\hat{a}_{\mathsf{ML}}$ is the value of a which maximises $F(m, a)$ with respect to a. Hence $\hat{a}_{\mathsf{ML}}$ is the solution for a of Equation 26. The limit $\underline{m}_{\mathsf{ML}} = 1.41$ is found by a numerical method.
2. Median: $\hat{a}_{\mathsf{med}}$ fixes the distribution of possible measured values such that the actual measured value is the *median*, hence $\int_{m'=0}^{m'=m} F(m', \hat{a}_{\mathsf{med}}).dm' = 0.5$. The threshold is $\underline{m}_{\mathsf{med}} = 1.18$, being the solution of $\int_{m'=0}^{m'=\underline{m}_{\mathsf{med}}} F(m', 0).dm' = 0.5$.
3. Serkowski's estimator: $\hat{a}_{\mathsf{Serk}}$ fixes the distribution of possible measured values such that the actual measured value is the *mean*, hence $\int_{m'=0}^{m'=\infty} m'.F(m', \hat{a}_{\mathsf{Serk}}).dm' = m$. The threshold is $\underline{m}_{\mathsf{Serk}} = 1.25 = \int_{m'=0}^{m'=\infty} m'.F(m', 0).dm'$.
4. Wardle & Kronberg's method: as defined above, the estimator, $\hat{a}_{\mathsf{WK}}$, is that which maximises $F(m, a)$ with respect to m (see Equation 27), and $\underline{m}_{\mathsf{WK}} = 1.00$.

Simmons & Stewart (1985) note that although widely used in the optical astronomy literature, Serkowski's estinator is not the best for either high or low polarizations; they find that the Wardle & Kronberg method commonly used by radio astronomers is best when $a \gtrsim 0.7$, i.e. when the underlying polarization is high and/or the measurement noise is very low. The Maximum Likelihood method, superior when $a \lesssim 0.7$ (i.e. in 'difficult' conditions of low polarization and/or high noise), appears to be unknown in the earlier literature.

In this paper, I have merely provided an interpolation scheme between the point estimators which they have shown to be appropriate to the 'easy' and 'difficult' measurement regimes. The construction of a confidence interval to estimate the error is actually independent of the choice of point estimator, although (as mentioned above) I believe that Simmons & Stewart's (1985, §3) unwillingness to 'accept sets of zero interval as confidence intervals' is unfounded, since physical intuition allows for the possibility of truly unpolarised sources (i.e. with identically zero polarizations), and their arbitrary method of avoiding zero-width intervals can be dispensed with.

11. Conclusion

The reduction of polarimetric data can seem a daunting task to the neophyte in the field. In this paper, I have attempted to bring together in one place the many recommendations made for the reduction and presentation of polarimetry, especially those of Simmons & Stewart (1985), and of Clarke *et al.* (1983). In addition, I have suggested that it is possible to develop the statistical technique used by Simmons & Stewart (1985) to obtain a simple probability that a measured object has non-zero underlying polarization. I have also suggested that there is a form of estimator for the overall degree of linear polarization which is more generally applicable than either the Maximum Likelihood or the Wardle & Kronberg (1974) estimators traditionally used, and which is especially relevant in cases where the measured data include degrees of polarization of order 0.7 times the estimated error.

Modern computer systems can estimate the noise on each individual mosaic of a sequence of images; this is useful information, and is not to be discarded in favour of a crude statistical analysis. A recurring theme in this paper has been the comparison of the errors estimated from propagating the known sky noise, and from applying sampling theory to the measured

intensities. Bearing this in mind, I have presented here a process for data reduction in the form of 18 rigorous steps and checks. The recipe might be used as the basis of an automated data reduction process, and I hope that it will be of particular use to the researcher – automated or otherwise – who is attempting polarimetry for the first time.

Acknowledgements

I would like to thank Steve Eales for his guidance with this project; Bob Thomson for his advice on statistics; and Jim Hough, Chris Packham, Mike Disney and Mike Edmunds for useful references. This work was funded by a PPARC postgraduate research student award.

References

Boas, Mary L. *Mathematical Methods in the Physical Sciences.* John Wiley & Sons, New York, second edition, 1983.

Clarke, G. M. and Cooke, D. *A Basic Course in Statistics.* Edward Arnold, second edition, 1983.

Clarke, D., Stewart, B. G., Schwarz, H. E., and Brooks, A. The statistical behaviour of normalized stokes parameters. [1] *Astronomy and Astrophysics*, 126:260–264, 1983.

di Serego Alighieri, S., Cimatti, A., and Fosbury, R. A. E. Imaging polarimetry of high-redshift radio galaxies. [2] *Astrophysical Journal*, 404:584–592, 1993.

Hecht, Eugene. *Optics.* Addison-Wesley, second edition, 1987.

Gehrels, Tom, editor. *Planets, Stars and Nebulae studied with Photopolarimetry*, chapter on Polarimetric Definitions (by D. Clarke), pages 45–53. The University of Arizona Press, 1974.

Leyshon, Gareth, and Eales, Stephen A. K-band polarimetry of seven high-redshift radio galaxies. *Monthly Notices of the Royal Astronomical Society*, in press, 1997. (E-print `astro-ph/9708085`.)

Mood, A. M., Graybill, F. A., and Boes, D. C. *Introduction to the Theory of Statistics.* McGraw-Hill, third edition, 1974.

NOAO. *On-Line Documentation for IRAF: APPHOT.PHOT.* On-line manual page for the PHOT command in the DIGIPHOT.APPHOT package of the IRAF data reduction system.

Serkowski, K. *Acta Astronomica*, 8:135, 1958.

Simmons, J. F. L., and Stewart, B. G. Point and interval estimation of the true unbiased degree of linear polarization in the presence of low signal-to-noise ratios. *Astronomy and Astrophysics*, 142:100–106, 1985.

Sterken, Chr., and Manfroid, J. *Astronomical Photometry: A Guide.* Kluwer Academic, 1992.

Vinokur, Marc. Optimisation dans la recherche d'une sinusoïde de période connue en presence de bruit. Application a la radioastronomie. *Annales d'Astrophysique*, 28:412, 1965.

Wardle, J. F. C., and Kronberg, P. P.. The linear polarization of quasi-stellar radio sources at 3.71 and 11.1 centimetres.[3] *Astrophysical Journal*, 194:249–255, 1974.

Notes

[1] In an attempt to use more consistent notation, my paper uses $\bar{s}$ for the arithmetic mean of a set of parameters, $\tilde{s}$ for a ratio of means, and $\hat{s}$ for the best (conservative) errors on certain quantities. Clarke *et al.*, however, use $\bar{s}$ for the ratio of non-normalized mean Stokes Parameters, and $\tilde{s}(1)$ for the arithmetic mean.

[2] My paper uses η for the instrumental angle which di Serego Alighieri *et al.* call ϕ.

[3] The Wardle & Kronberg paper reproduces Vinokur's equation (my Equation 38) but omits the factor 'sign(x)' from Equation 39 on the grounds (Wardle, private communication) that the probability of x falling in the domain $x < 0$ is negligibly small.